

It will be some time, however, before operators can specify searches in natural language, and there is still no equivalent to the SQL structured data query language suitable for use with unstructured data. Although full-text retrieval technology has come a long way from the days of disk farms and dedicated mainframes, further breakthroughs are still needed before every publisher's dream of HAL-like access to oceans of data can be realized.

### The hype about hypertext

Thanks mainly to Apple's hype about its Hypercard product, virtually everyone in the industry has been paying homage to hypertext recently. What isn't well known is that hypertext has more divergent sects than Protestants or Muslims, each claiming with an almost religious fervor to be the One True Voice on the subject.

In this year's seminar we decided to go directly to the source of most of the original hype in hypertext, Ted Nelson, and juxtaposed with him a somewhat cynical viewpoint on the matter from Bob Marsh of Plexus, who has had the real-world experience of implementing hypertext access to an enormous on-line database his company is assembling for the University of Rochester.

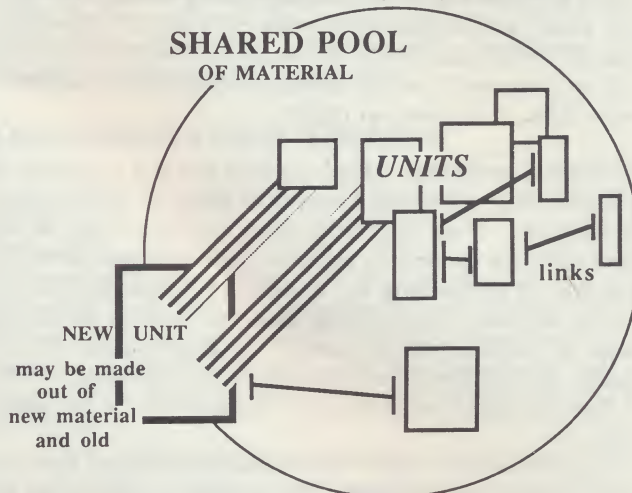
**Paper is not dead; however.** . . . Ted Nelson began his sermon by reassuring the assembly that paper is not dead—that it has at least two more decades as a viable medium. Nelson is confident, however, that the cost of gasoline and the shrinking forests will push us to do most of our publishing electronically in the future.

**The ever-expanding docuverse.** According to Nelson, hypertext is a document "that is not constrained by linear sequences." Nelson's Xanadu Project is an attempt to provide an "organic," evolving, ongoing database that begins on day 1 and continues on indefinitely to incorporate any number of nodes, distributed worldwide (and in space). Each contributor will be able to add his document to a shared pool of material, which makes up the "docuverse" (a coinage that became the Seminar's most-quoted-out-of-context term by the end of the week).

The key question in looking at how we move up to a system in which all publishing is electronic is one of scaling up. If access time increases linearly with the number of documents, it will be impossible when the order of magnitude exceeds thousands. To support millions/billions of documents, we need to assure that access time grows (at most) as the square root of system size, preferably as the logarithm of size. Xanadu is being designed to assure that performance degradation gets no worse than square-root rates.

Xanadu has some nice properties. A write-once system guarantees that nothing is lost or corrupted (ever, since Nelson expects storage technology enhancements to keep pace with Xanadu system growth). Quotes are always accurate, in context, and correctly attributed. The original context remains available. Originals are quoted by placing a pointer in the commentary; the system records the fact of the quotation and could therefore automatically collect royalties on every access of copyrighted material (but this requires that every document have a royalty price—no one could refuse to

## XANALOGICAL STORAGE



**Project Xanadu quotations.** In an electronic publication, quoted material is not repeated, but merely referenced by its computer address. When the publication is read, the system finds all references and delivers their original contents to the reader.

license material). All formatting of documents resides in links (which are documents) rather than being embedded in the content. In this regard, the Xanadu system is not much different from other systems that separate form from content.

Xanadu is currently running on a Sun-3/60 workstation and utilizing a little over 130 KB of memory. Needless to say, the system will require enormous quantities of storage.

### Hypertext without the hype

We couldn't begin to describe the data structure and access conventions that we are told will keep the Xanadu system hopping along for years, well into the docuverse. In any event, the problems associated with creating a database from scratch are altogether different from attempting to catalog, input, and structure a pre-existing paper database.

Such was the case when Plexus implemented a system for the Syracuse University Library's Adult Education collection. This consisted of about 750,000 pages and 15,000 slides plus tapes, records and films. The University had received a grant from the Kellogg Foundation to design a system to handle these requirements, using the latest technology, over a five-year period.

Plexus is assembling a system that uses Ethernet as the backbone, AT clones with 1600×1200 displays running Windows 2.0 as the user workstations, a Palantir OCR server, scanners, etc. The main file server is a Plexus computer with a jukebox of optical disks, running Unix and the Informix relational database. A worldwide network is contemplated, using TCP/IP for computer data and facsimile for images.

Partially because of this experience, and partially because he is a careful, rational thinker with extensive experience and knowledge in the computer industry, we invited Bob Marsh, chairman of Plexus, to speak as a counter-balance to Nelson.

**Plexus model.** In the Syracuse system "items" are essentially file folders with tagged text and hypertext linkage pointers. They contain "components" that are files of text, images, etc.

This leads to the specter of several different PostScript font libraries. All would likely share compatible versions of the 35 base fonts. Beyond that, they would provide different fonts with different and incompatible font metrics.

### Compugraphic's Ned Bunnell

CG is a traditional typesetter/font supplier which now finds itself in two separate but related businesses: output devices and type fonts. It bills itself as the industry's largest type company with 300 employees, 1700 faces with a total of 1 million unique characters, and a production of 100 new fonts per year.

CG's emphasis as an independent supplier of type fonts is on its Intellifont program. Like Bitstream FontWare, this uses special "intelligent" font outlines and scaling routines to produce screen and printer font bit maps. Intellifont 2.0 uses a line and arc outline data format (CG previously used only straight lines) and facilities for character manipulation. There are now 145 Intellifonts, including the infamous 35 tuned to match Adobe metrics. Like Bitstream, CG believes that Postscript clone font libraries will match font metrics only for these base type faces.

CG has a deal with Hewlett-Packard under which it will supply fonts for PCL printers. This could make the CG library the *de facto* standard for PCL printers. But CG also expects to play in the Postscript market (see preceding paragraph), and in the Apple QuickDraw printer market. The announcement of CG fonts for Apple's new LaserWriter IIsc QuickDraw printer was one of the most interesting developments of the session. The so badly needs well-designed printer fonts (with corresponding screen fonts). CG could be the first player in a growing market.

### The Company's Mike Parker

Mike Parker of The Company spoke, representing his firm and his client, URW. URW had several interesting developments to talk about. It promised to have Bezier outline fonts (required for Postscript engines) by April 1, with automatic insertion of scaling hints by computer program (rather than by hand as everyone currently does it) by the end of the year. URW also launched MacIIkarus, a version of the Ikarus outline type generation software running on a Macintosh. Parker sees this as appealing to trade type houses, design firms and others who would love to be able to create their own specially modified typefaces. This is real democratization of the typesetting process: putting type design into the hands of the user!

### Xiphias's Peter Black

Peter Black's Xiphias is taking a very different tack from the other vendors on the panel. Rather than optimizing fonts for printer output, then trying to make passable screen versions of these faces, Xiphias is concentrating on applications in which the screen display and/or output on crude devices (such as 24-pin dot matrix printers) is crucial. In a nutshell, Xiphias specializes in fonts for "anything that's dumb." For such devices, it expects that there will be

Louis Villalobos, president of Conographic, was his usual controversial self. First, he claimed that the Conographic font-scaling technology, which is based on second-degree parametric equations (conic sections), can eliminate the need for special hints—that it can generate good-looking type at 300 dpi from any outline font. Indeed, Conographic is currently licensing fonts from several font houses, as well as designing its own Conofonts. This technology is for sale to all comers.

None of the panelists responded to this subject area, though most of them—Compugraphic, Bitstream, Adobe, Linotype, URW—have invested heavily in hint-based technologies. Whether hints are really necessary is a testable proposition, but we have not yet seen the proof in a real product. Conographic's show demos have thus far been dumping canned files from disk to printer.

Second, Villalobos raised the spectre of the Rockwell-MGD patents now owned by Information International. Conographic is convinced that its conic section font technology does not violate the Rockwell patents, and it said that it would offer its customers an indemnity against patent-infringement lawsuits. (We doubt that Conographic's pockets are deep enough to pay multi-million-dollar settlements, the way Compugraphic and Linotype did, which is the other way of viewing Villalobos's statement.)

Rockwell patents. The issue of patent indemnity raised—but did little to settle—the question of whether the Rockwell patents would hold up in court if anyone ever let things get that far. Several years ago, first Linotype and more recently Compugraphic chose to settle quietly rather than face a legal test. They both now presumably have a vested interest in the validity of the patents. The patents would constitute an expensive barrier to entry for their competitors. Mike Parker, however, averred that the patents make excessively broad claims, that there is prior art covering the same material, and thus that the patents would probably not withstand the scrutiny of a court. Left unspoken was the possibility that Adobe, Bitstream, or URW might themselves have to decide whether to settle or fight at some point in the future.

Whose Times? This discussion led directly to the question of protecting font designs, or more specifically, "Who owns Times Roman?" Both Linotype and Monotype have claims in this area. Linotype has a trademark in the U.S. on the name, and has licensed the design to Adobe Systems and Imagen. Monotype claims that it was granted sole licensing rights by the owner of the design, *The Times* of London. Once again, Mike Parker gave a definitive history. In his version of the story, *The Times* licensed both Linotype and Monotype to make and sell the font family (in those days as hot-metal matrices), but additionally gave Monotype the power to re-license manufacturing rights. By implication, then, Linotype did not receive that power.